AD NUMBER

AD417243

NEW LIMITATION CHANGE

TO

Approved for public release, distribution unlimited

FROM

Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; AUG 1963. Other requests shall be referred to Office of Naval Research, Washington, DC 20350. Availablity: Document only can be viewed at DTIC, mostly illegible.

AUTHORITY

Avail st'mt iaw fmq ltr, 24 Sep 1991

UNCLASSIFIED

AD 417243

DEFENSE DOCUMENTATION CENTER

FOR

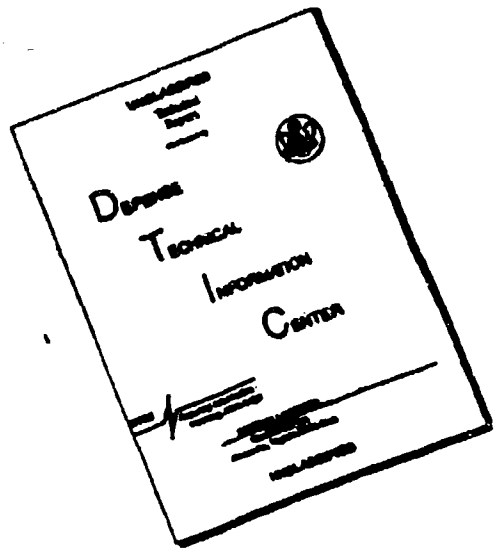
SCIENTIFIC AND TECHNICAL INFORMATION

CAMERON STATION, ALEXANDRIA, VIRGINIA

CODE 22

UNCLASSIFIED

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

SOME RELATIONS BETWEEN DIGITIZING PARAMETERS AND CALCULATED STATISTICS OF A WAVEFORM

by
ROBERT McAULAY

15 August 1963

Technical Report No. 18
Contract Nonr 1834(02)
ONR Project No. N — 371-161

and

Technical Report No. 1
Contract Nobsr-89229
Bureau of Ships

Supported by
OFFICE OF NAVAL RESEARCH
and
BUREAU OF SHIPS

RADIOLOCATION RESEARCH LABORATORY
REPORT NO. R.R.L.-208
DEPARTMENT OF ELECTRICAL ENGINEERING
ENGINEERING EXPERIMENT STATION
UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS

SOME RELATIONS BETWEEN DIGITIZING PARAMETERS AND CALCULATED STATISTICS OF A WAVEFORM

by

Robert McAulay

15 August 1963

Technical Report No. 18
Contract Nonr 1834(02)
ONR Project No. NR 371-161

and

Technical Report No. 1
Contract Nobsr-89229
Bureau of Ships

Supported by
OFFICE OF NAVAL RESEARCH
and
BUREAU OF SHIPS

Radiolocation Research Laboratory
Report No. R.R.L. -208

Department of Electrical Engineering
Engineering Experiment Station
University of Illinois
Urbana Illinois

ACKNOWLEDGEMENT

The author wishes to thank Dr. Robert Smith for his helpful suggestions and guidance which he generously gave throughout all stages of this work. His interest and encouragement were also appreciated. The author also thanks Prof. A. D. Bailey for his comments on the final draft.

This research was carried out under contracts sponsored by NONR 1834(02) and NOBSR 89229.

CONTENTS

ILLUSTRATIONS

1. STATEMENT OF THE PROBLEM

One of the problems which arises when a computer is utilized in data-analysis systems is the conversion of the incoming data to a form which can be used as input to the computer. When a digital computer is used in the system, this conversion is usually made by some sort of analog-to-digital converter which transforms the incoming electrical signal into a new signal of a more useable form. For instance, this new signal can be fed directly into the computer through some type of line coupler, or be used to punch a set of coded numbers on paper tape or cards which can be input to the computer at a later time.

In many cases the analysis to be performed is statistical in nature in as much as the computed output is the value of certain statistics which enables the experimenter to make decisions about the phenomenon being observed. These decisions should always be weighed by acknowledging the errors which are inherent in a system of this type. In this case the error is introduced by the hardware used in the conversion process which transforms the data from analog to digital form.

The reason for this error is the following: Initially the incoming electrical signal is sampled in time and a new waveform is produced which has discrete amplitudes at discrete time intervals. The magnitude of each of these amplitudes is then "rounded off" to some predetermined amplitude level which is nearest to the value of the sampled amplitude. This is referred to as the quantization process and the predetermined amplitude levels are called quantization levels. The combined process, sampling in time and quantizing the resulting amplitude is called the digitizing process and results in a set of discrete data points. The computer uses this latter set of points to calculate

certain statistical quantities which represent, in some sense, the original electrical signal and therefore the physical phenomenon being studied.

The question which arises is whether or not the computed statistics are in fact representative of the true values of the statistics which would be obtained if it were possible to compute them for the incoming electrical signal.

By simulating the digitizing process on a digital computer and operating on simple known waveforms this paper attempts to develop a technique which might prove useful in investigating these problems.

2. SIMPLIFICATION OF THE PROBLEM

It has been pointed out in the previous section that the signal transformation of primary interest is the digitizing process. In particular it is desirable to investigate the effect which the analog-to-digital converter has in perturbing the values of certain statistics from their true values. One way of attacking this problem is an experimental one whereby the analog of the actual sampling and quantizing process is developed as a computer program. This program can then be used to operate on the mathematical analog of the waveform under consideration to produce a new set of data points which represents the quantized amplitudes at successive sample times. Using this data the sample statistics can be calculated and compared with the "true" values which, in this case, can be calculated directly using the mathematical expression for the incoming waveform. This procedure can be repeated for many combinations of sampling rates and quantizing levels until some agreement or lack of agreement between the two sets of results can be determined.

Figures 2.1a and 2.1b show the waveforms which result when the digitizing process is applied to a sine wave superimposed on a steady value. It is fairly obvious from even these simple sketches that the digitized waveform having a larger number of samples and a larger number of quantization levels "resembles" more closely the original sine wave. It seems reasonable to say that the statistics in this latter case might have values nearer to the true values calculated on the basis of the original waveform.

Throughout the remainder of this paper only the sine wave will be considered. This may seem to be a gross simplification to make but if any other periodic waveform is to be used as an input signal it can be Fourier analyzed into a sum of sine waves. Therefore before any statement can be made regarding the effects of the digitizing process on more general waveforms, it is necessary to examine the case of the single sine wave in detail. Since the Fourier series is a sum

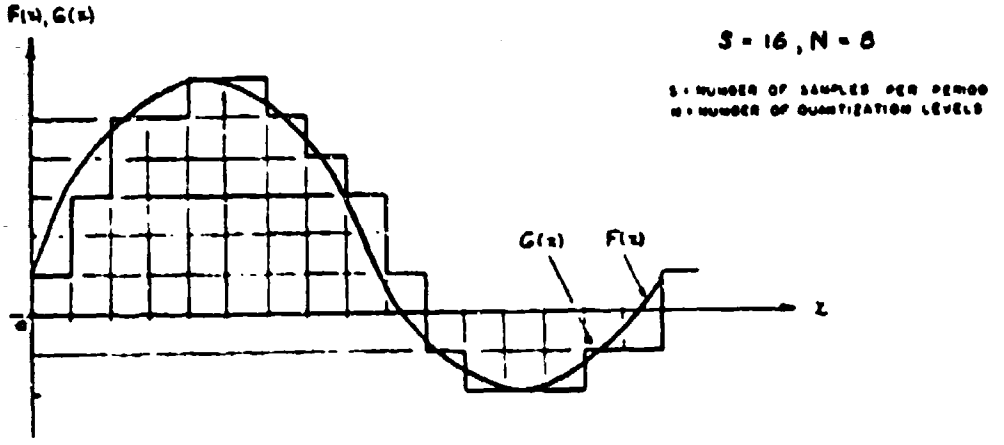


Figure 2.1a

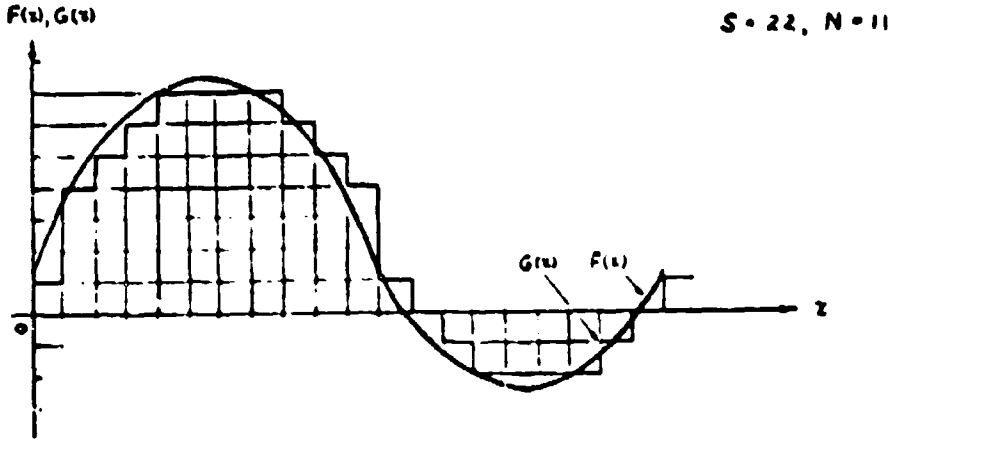


Figure 2.1b

F(x)= A sine(x+b)+B, A, B, b constant
Approximate waveforms obtained using the quantizing/sampling Technique

of sine waves, a linear operation, the effects on the statistics for the original periodic waveform can be expressed as the sum of the effects obtained for each component sine wave.

2.1 The Sampling Process

The sampling process is nothing more than a quantization of the argument of the function which represents the incoming waveform, in this case the sine wave. This process is subject to the constraints imposed by the well-known Sampling Theorem[1] which gives the relation between the rate at which a signal varies and the number of sampling pulses needed to reconstruct it. In particular the Sampling Theorem says that the power spectra of the sampled and unsampled waveforms are the same provided certain constraints on the sampling frequency are satisfied. When the power spectra are the same then the first two statistical moments are the same. Since the present work makes use of only the first two statistical moments, the constraints imposed on the sampling frequency by the Sampling Theorem will have to be satisfied.

The reason for considering only the first and second statistical moments is the following. If the incoming waveform represents a voltage or a current then the first moment, called the simple mean, represents the DC component of that voltage or current. The second moment represents the total power which that voltage or current would dissipate in a 1 ohm resistor. The square root of this quantity is the RMS value of that voltage or current. Subtracting the square of the simple mean from the second moment gives the alternating power which that waveform would dissipate in a 1 ohm resistor. The square root of this quantity gives the AC component of that waveform. Consequently since electrical signals are usually characterized by their average, peak, and RMS values, only the first and second moments need be studied.

To simplify the analysis which follows, the incoming waveform is considered to have zero mean value. That is, the first moment is assumed to be zero. Therefore only the sinusoidal component will be studied.

After the waveform has been sampled, each amplitude is perturbed by the quantization process and this in turn will perturb the value of the sample statistic. Consequently even if the constraint on the sampling frequency is satisfied, as required by the Sampling Theorem, this will no longer be a sufficient condition to guarantee the equivalence of the sample values and the true values of the statistics. However, the constraint does provide a lower bound on the sampling frequency since, if there were no quantization, the theorem would be satisfied and the values of the statistics would be the same.

If f_m is the highest frequency component or the bandwidth of the incoming signal and if f_s is the rate at which the samples are taken, then the Sampling Theorem requires that

$$f_s \geq 2f_m \tag{2.1}$$

That is, at least $2f_m$ uniformly spaced samples are required every second to reproduce the statistics of the sampled waveform, where the amplitudes at each sample time have not been perturbed by the quantization process. Consequently it will be necessary to specify the bandwidth of the incoming signal so that the lower bound on the sampling rate can be determined. This is a particularly simple matter for the case under consideration since the sine wave is defined by a single frequency. Sampling must occur at least twice that frequency resulting in at least two samples per period. Equation (2.1) can be rewritten as

$$f_s = S f_m \qquad \text{where } S \geq 2.$$

The period between sampling pulses is

$$T_s = \frac{1}{f_s} = \frac{1}{Sf_m}$$

The argument of the function is x where

$$x = \omega_m t \quad \text{and} \quad \Delta x = \omega_m \Delta t = 2\pi f_m \Delta t$$

The time between samples is Δt and therefore

$$\Delta t = T_s = \frac{1}{Sf_m}$$

Therefore the sampled argument is

$$\Delta x = \frac{2\pi}{S} \quad \text{where } S \geq 2$$

and

$$x = (i \Delta x) = \frac{2\pi i}{S} \quad i = 1, 2 \ldots S \qquad (2.2)$$

The quantity S, the number of samples per period is one of the parameters of primary interest.

Since the waveforms which will be considered throughout the rest of the paper are periodic, they can always be defined in terms of the argument x. To simulate the sampling process on the computer, therefore, it is necessary to store the mathematical description of f(x), for all x between 0 and 2π, and call out those

amplitudes given by $f(2\pi i/S)$ where the number of samples to be taken (i.e. S) has been specified at the outset and i is allowed to index through values from 1 to S. The computer program which was developed was a direct application of these two steps.

2.2 The Quantization Process

The most general equi-interval quantizer has the transfer characteristics shown in Figure 2.2. This quantizer can be decomposed into a system containing shifts, gain factors, and a unit quantizer ($\Delta V = 1$, $a = 1$, $r = 1$) from which a mathematical expression relating the input and output waveforms can be derived. This has been done in the paper by D. G. Watts[2]. For the present it is sufficient to utilize only the following result: If f(x) is the input waveform and q(x) is the quantized output waveform, quantization occurs when

$$(n - \tfrac{1}{2})\,\Delta V < \left|f(x)\right| < (n + \tfrac{1}{2})\,\Delta V \qquad n = 0, 1, 2, \ldots N \qquad (2.3)$$

where N is the number of quantization levels on each side of the zero amplitude level. The resulting output is

$$q(x) = n\Delta V \qquad \text{if} \quad f(x) > 0 \qquad (2.4)$$

$$q(x) = -n\Delta V \qquad \text{if} \quad f(x) < 0 \qquad (2.5)$$

Figure 2.3 illustrates the manner in which the quantization levels are assigned for the simple case when $N = 4$. The quantization interval ΔV is 1/N where N is actually only $\frac{1}{2}$ of the total number of quantization levels used since it refers

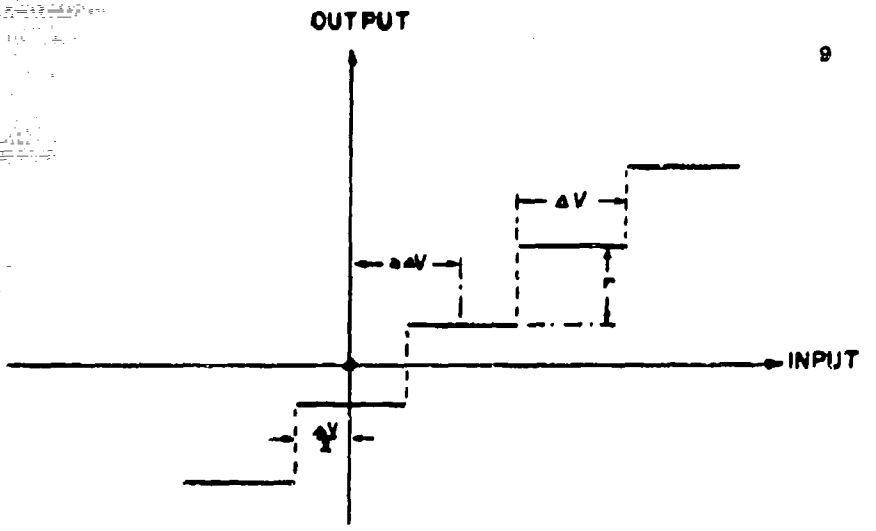


Figure 2.2 Transfer characteristics of general equi-interval quantizer

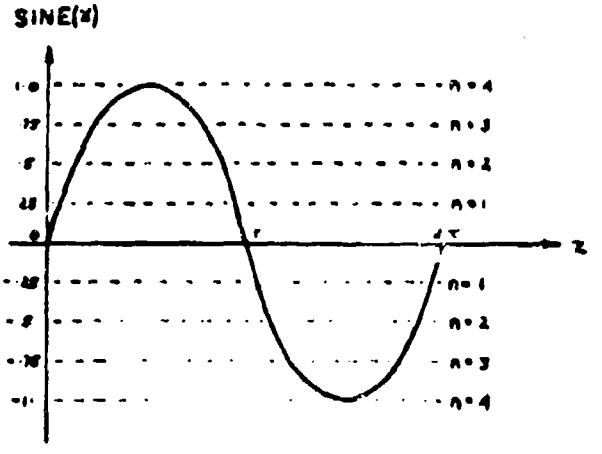


Figure 2.3 Bilateral Quantization - $\left\{\begin{matrix}4 \text{ levels positive} \\ 4 \text{ levels negative}\end{matrix}\right\} N = 4$

to the number of levels on each side of the zero voltage level.

Using Equations (2.3), (2.4), and (2.5) it was possible to simulate the quantization process on the computer. To obtain the analog of the entire digitizing process (i.e. sampling and quantizing), it remains only to combine the two processes just considered. In fact, the only revision which has to be made is in Equation (2.3) where instead of quantizing f(x) for all x, only f(2πi/S) need be quantized for the respective sample arguments which are obtained when i indexes from 1 to S. A sample calculation will be considered next in order to illustrate the main points of the preceding discussion.

2.3 Sample Calculation

$$f(x) = \sin(x) \quad 0 \leq x < 2\pi$$

Suppose that f(x) is sampled at a rate of 10 sample per period, (i.e. S = 10) and a total of 20 quantization levels are used, (i.e. N = 10). The increment of the argument is

$$\Delta x = 2\pi/10 = 0.157 \text{ radians}$$

Therefore at each sample time

$$x_i = i\Delta x = (0.157i) \text{ radians for } i = 1, 2, 3, \ldots 10$$

The quantizing interval is

$$\Delta V = 1/N = 1/10 = 0.1$$

The analog of the sampling and quantizing process is therefore given by:

$$0.1(n - \tfrac{1}{2}) \leq \left|\sin(0.157t)\right| < 0.1(n + \tfrac{1}{2}) \qquad n = 0, 1, \ldots 10$$

and

$$q(0.157t) = +0.1n \quad \text{for } \sin(0.157t) > 0$$

$$q(0.157t) = -0.1n \quad \text{for } \sin(0.157t) < 0$$

The resulting digitized sine wave is illustrated in Figure 2.4. The amplitudes obtained at each sample time and the square of each of these amplitudes were summed over all sample times and the following results were obtained

$$\sum_{t=1}^{40} q(0.157t) = 0$$

$$\sum_{t=1}^{40} q^2(0.157t) = 0.4945$$

Using these results the sample statistics were calculated as follows

$$\text{sample mean} = \frac{1}{40} \sum_{t=1}^{40} q(0.157t) = 0$$

$$\text{sample variance} = \frac{1}{40} \sum_{t=1}^{40} q^2(0.157t) - (\text{sample mean})^2 = 0.4945$$

$$\text{sample standard deviation} = \sqrt{0.4945} = 0.703$$

Using the probability function for the sine-wave process the true mean and variance can be calculated. A measure of how well the sample values compare with these true values can then be made.

The probability distribution function for the sine-wave process is (cf. Brunck[3]):

$$P[x(t) \leq X] \quad \begin{cases} 1 & X > +1 \\ \frac{1}{2} - \frac{1}{\pi} \sin^{-1} X & -1 \leq X \leq +1 \\ 0 & X < -1 \end{cases}$$

and the probability density function is

$$p[x(t)] \quad \begin{cases} 0 & |X| > 1 \\ \frac{1}{\pi \sqrt{1 - X^2}} & -1 < X < +1 \end{cases}$$

The graphs of these functions are illustrated in Figure 2.5. Using these relations the true mean is

$$E[x(t)] = \int_{-1}^{+1} X \frac{1}{\pi \sqrt{1 - X^2}} \, dX = 0$$

and the second moment is

$$E[x^2(t)] = \int_{-1}^{+1} X^2 \frac{1}{\pi \sqrt{1 - X^2}} \, dX = 0.5$$

Therefore the variance is

$$E[x^2(t)] - E[x(t)]^2 = 0.5$$

and this results in a standard deviation of 0.707

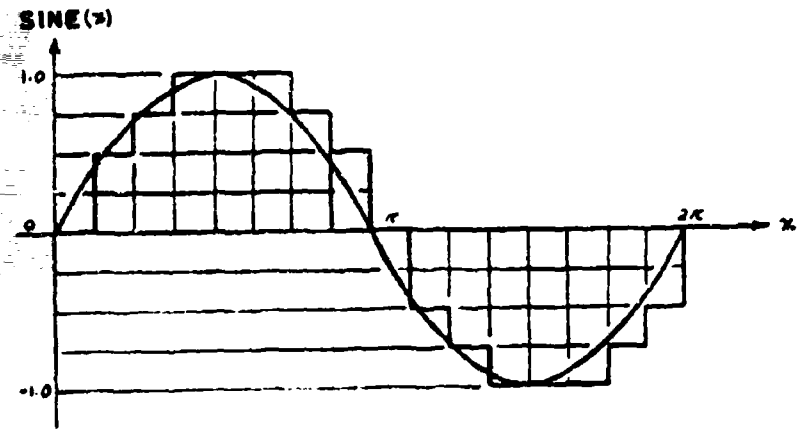


Figure 2.4 A sine wave and its digitized equivalent

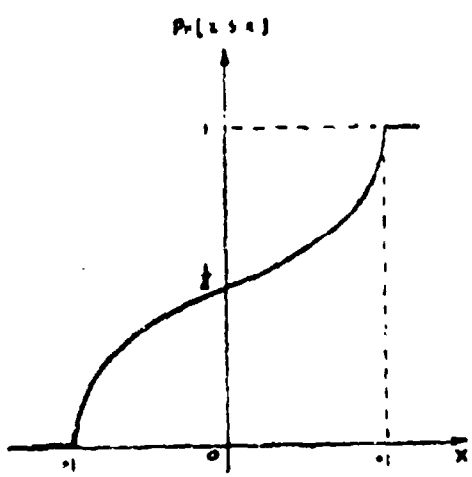
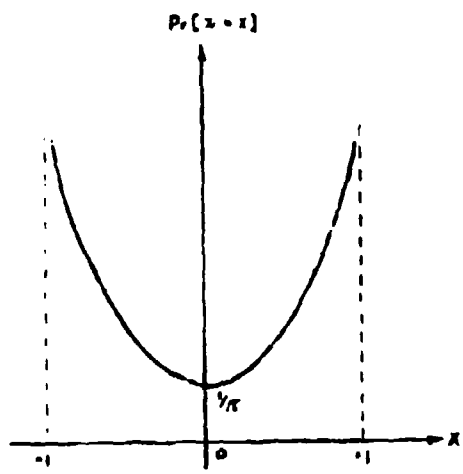


Figure 2.5 Distribution Function



Density Function

If a closer look at the digitizing process for the sine wave is taken, it will be observed that when the sine wave has zero phase it always has zero sample mean for all values of S. For every positive amplitude there is a corresponding negative amplitude of the same value and the resulting sum will therefore always be zero. If the sine wave is shifted with respect to the t = 0 axis, then the sample mean will be zero only for even values of S. In both cases, when the sampled and quantized amplitudes are squared, cancellation cannot occur and the sample variance has a value which in general is not equal to the true value, for all S and any phase. Because zero phase has been assumed for the sine wave under consideration it will be necessary only to examine the errors in the sample variance since the mean is zero for all S as pointed out above. This error is best described in terms of the sample standard deviation since it has dimensions which are more meaningful physically. If the sine wave is a voltage, for example, the standard deviation will also have the dimensions of volts.

In this case the percentage error in the standard deviation is

$$\frac{0.703 - 0.707}{0.707} \cdot 100\% = -0.56\%$$

When a more general periodic waveform is considered which has more than one frequency component, it is best to redefine the previous result in terms of the variance. As has already been pointed out the variance corresponds to the AC power dissipated in a 1 ohm resistor when the process x(t) represents a voltage or a current. Since the total AC power of the periodic waveform is the sum of the powers of each frequency component, the total variance will be the sum of the variances of each component sine wave. Making use of this idea it is possible to extend the results obtained for the single sine wave case to predict the results for slightly more general cases.

Although the computation involved in obtaining the above results was quite simple conceptually, it was very laborious and time consuming. Furthermore, the results were based on only one value of S and one value of N, while it is really necessary to know the behavior for many values of S and N. It was possible to obtain a rather large quantity of data in a relatively short time by programming the computer to simulate the digitizing process in the manner just described.

The section which follows deals with the analysis of this data.

3. EXPERIMENTAL PROCEDURES AND RESULTS

Making use of the computer analog of the digitizing process developed in the preceding section it was possible to obtain the difference between the values of the standard deviation calculated for the unperturbed waveform and the corresponding quantized waveform as a function of the number of quantizing levels over a wide range of sampling rates. It was found that these differences (expressed as percentage errors with respect to the true value) tended to become smaller as the number of quantization levels increased. Figure 3.1 is a typical plot showing the errors which were obtained when the sampling rate was fixed at 20 samples per period and the number of quantization levels ranged from 1 to 256. In this form it is difficult to make efficient use of the data.

For the most part, this paper is concerned with the matter of describing any tendency which the errors in the standard deviation might demonstrate with respect to the values of S and N. Therefore a restricting criterion, which can be used to reduce the quantity of data which has to be analyzed, is to order the maximum errors with respect to an increasing value of N. In reference to Figure 3.1, for example, the largest error obtained over the entire range of N occurs at N = 6. By definition of the ordering process, the smaller errors which are obtained for N less than 6 will be neglected. The remaining errors over the range N greater than 6 are scanned until the next largest error is found. In this case this occurs at N = 7. In general the errors between N = 6 and N = 7 would be neglected, again by definition of the ordering procedure. (Unfortunately this was a trivial case because no errors were obtained in this interval). The range for N greater than 7 is then scanned for the next largest error which, in this case, is obtained at N = 11. As before the smaller errors in the interval between N = 7 and N = 11 are neglected. By repeating this scanning procedure

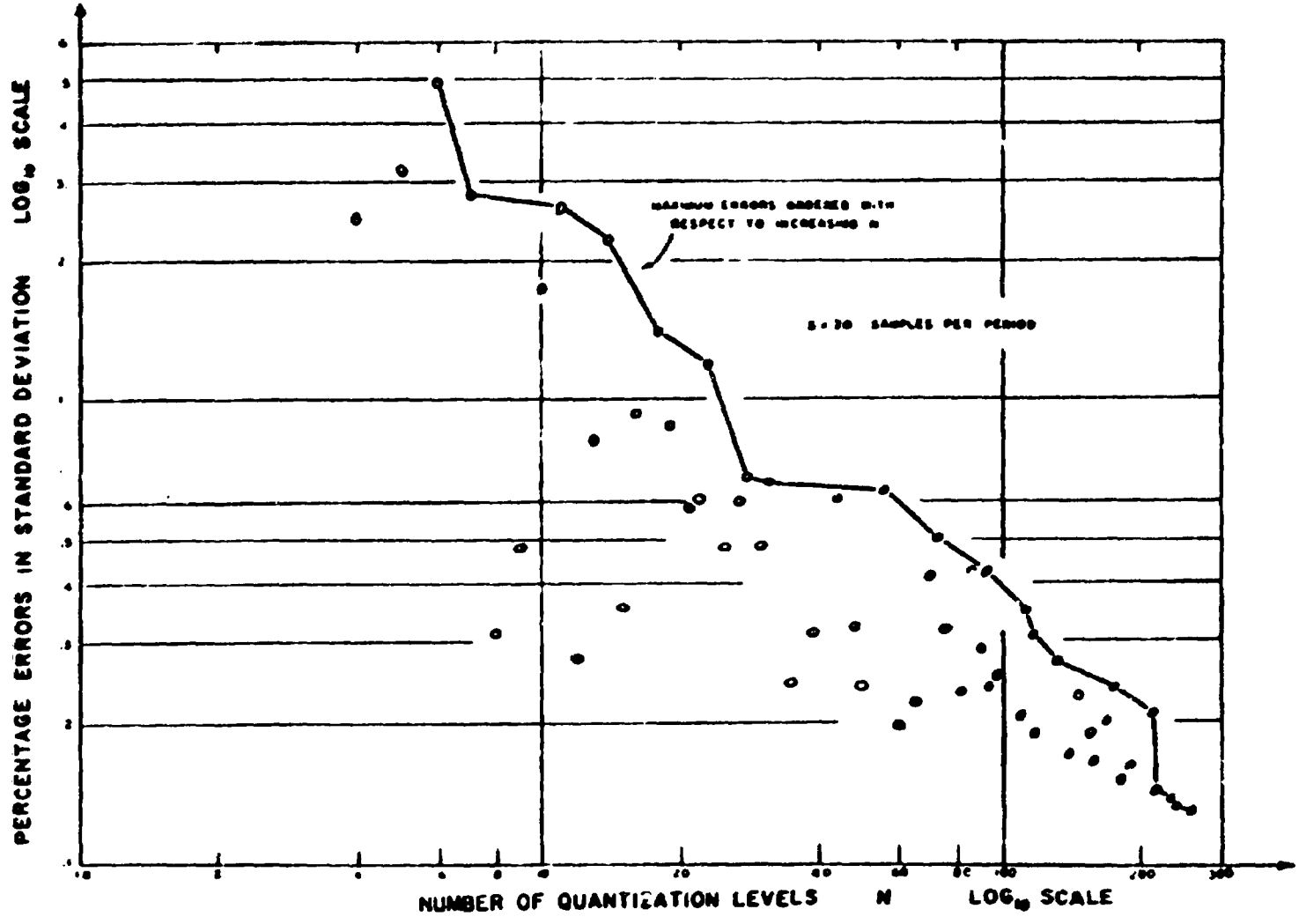

PERCENTAGE ERRORS IN STANDARD DEVIATION LOG10 SCALE
NUMBER OF QUANTIZATION LEVELS N LOG10 SCALE

Figure 3.1

for the remaining values of N the curve drawn in Figure 3.1 is produced. By considering only the errors given by this curve it is a certainty that for each value of N the errors are either less than or equal to the values given by the ordinate of that curve. Obviously any analysis based on this curve will be based on "worst-case" results.

It should be noted that this technique certainly does not make the most efficient use of the data (in the statistical sense), but for this preliminary investigation it does reduce the quantity of data to a more manageable level.

Another curve was drawn in the same way giving the "worst-case" results when 100 samples per period were taken (i.e. S = 100). Figure 3.2 illustrates this curve as well as that plotted in Figure 3.1 for S = 20. The straight lines represent the best fit linear approximation to the original curves. The significant observations which can be made on the basis of these curves are the following:

1. The curves are approximately linear in terms of $\log_{10}$ values of E and N, where E is the percentage error in the standard deviation of the digitized waveform and N is the number of quantizing levels. This suggests that the relationship between E and N is a rectangular hyperbola. The straight-line $\log_{10}$ plot is given by

$$\log_{10} E = -m \log_{10} N + b \qquad (3.1)$$

or

$$E = B N^{-m} \qquad (3.2)$$

2. As the sampling rate increased the E vs N curves are displaced in the direction of decreased error.

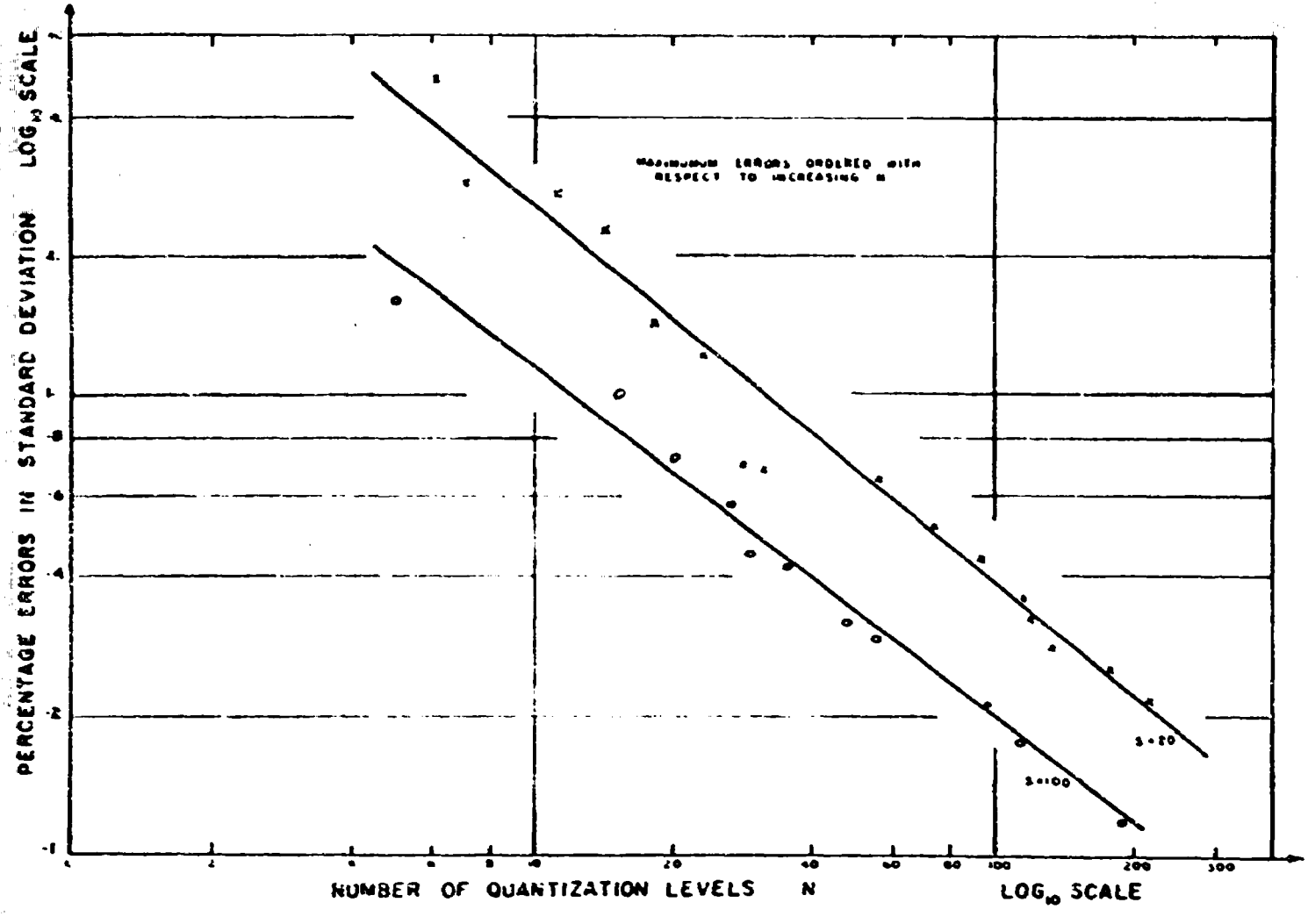

PERCENTAGE ERRORS IN STANDARD DEVIATION LOG10 SCALE
NUMBER OF QUANTIZATION LEVELS N
LOG10 SCALE
S=20
S=100

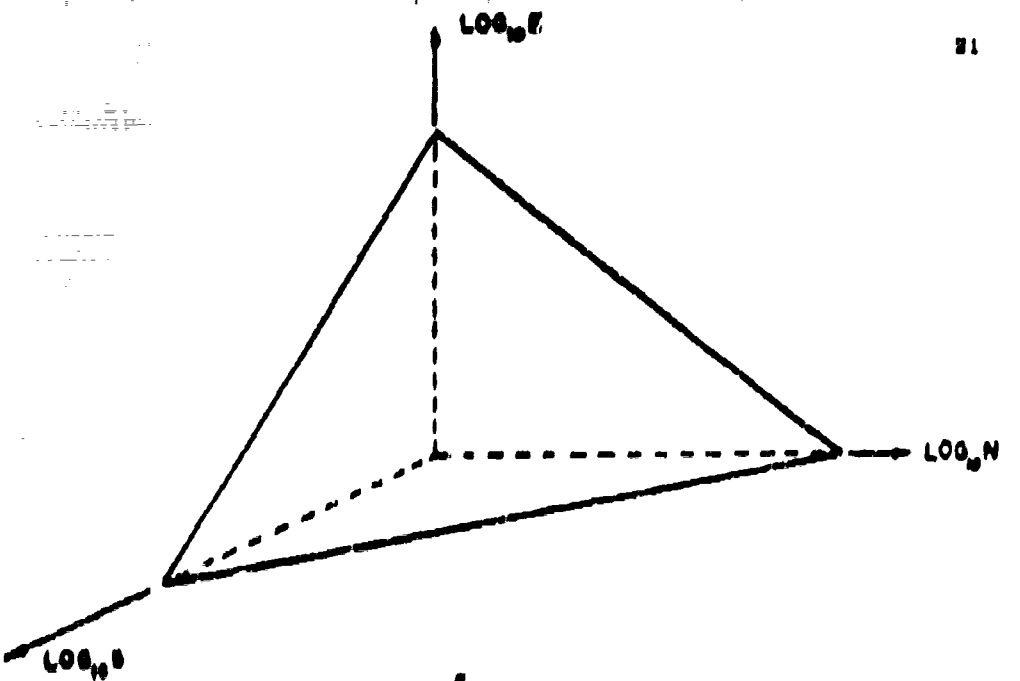
Figure 3.2

These results suggest a three dimensional surface of the form shown in Figure 3.3. Equations (3.1) and (3.2) represent the S = constant cross sections. If this surface can be described analytically the composite E-N-S relationship will be known completely. This would undoubtedly be a very useful result. The following paragraph describes how this analytical expression was obtained.

At first a rather large matrix was constructed having S rows and N columns. Each element of this matrix was the error obtained for a particular value of S and N which correspond to the subscripts of that element. The points plotted in Figure 3.1 produce the elements of row 20 and columns 1 to 256. Using a technique similar to that used when S was one fixed value, the maximum errors are again ordered but now with respect to both increasing N and increasing S values. By using the computer to generate the errors for each value of S and N, as described in Chapter Two, this matrix was constructed fairly easily and quickly.

$$\begin{bmatrix} \epsilon_{11} & \epsilon_{12} & \cdots & \epsilon_{1n} \\ \epsilon_{21} & \epsilon_{22} & \cdots & \epsilon_{2n} \\ & & & \\ \epsilon_{s1} & \epsilon_{s2} & \cdots & \epsilon_{sn} \end{bmatrix}$$

After the last element ϵ_{sn} had been computed and stored, the search for the largest error was initiated. This was done by comparing the magnitudes of the elements ϵ_{ij} and $\epsilon_{i,j-1}$, $j = 1,N$, and storing the larger number. The entire matrix was scanned in this way and the largest value was retained along with its subscripts. Suppose for example, that ϵ_{24} was the largest element. Then the



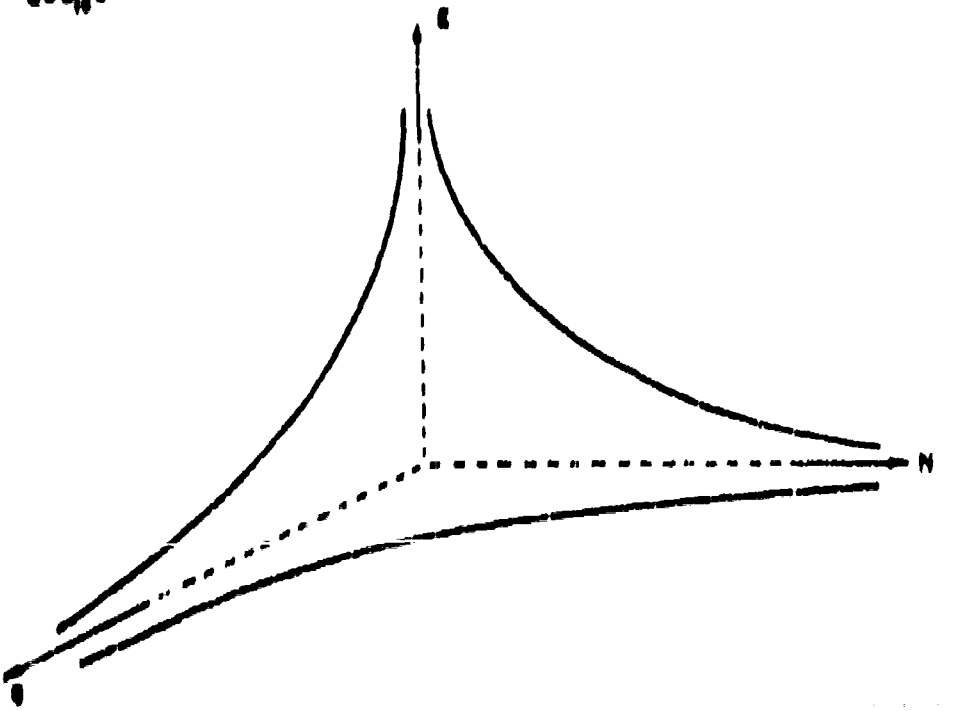


Figure 4.3 Nomogram representing the B-N-E relationship

magnitude of ϵ_{24} was stored as the first element in a new matrix with its new position flagged by S = 2 and N = 4. As before, the preceding elements $\epsilon_{11} \cdots \epsilon_{1n}$, ϵ_{21}, ϵ_{22}, ϵ_{23} were ignored. The computer then indexed to the next element (i.e. ϵ_{25}) and repeated the search procedure until the next largest error was obtained for S = 2 and N = 5. This element was stored as the second element of the new matrix along with the values of its subscripts as before (viz. S and N) This search and storage procedure was repeated until all the elements of the original matrix were exhausted. As a result, the new matrix gives the points on the "maximum-error surface" with respect to increasing values of S and N. It is therefore a certainty that for any S and N within the specified range the error which can be expected for those coordinates is less than or equal to the value given by the point on the surface. As a result any analysis based on this surface gives worst-case results.

The next problem was to obtain the equation of the three dimensional surface from the data contained in the matrix of the ordered maximum errors. The curves drawn in Figure 3.2 have demonstrated a fairly-well defined relation between E and N as given by Equations (3.1) and (3.2). Because of the tendency to displace the curves in the direction of decreased error when the sampling rate was increased, it is not unreasonable to assume a similar relations between E and S as was assumed for E and N. In three dimensions this results in the equation

$$\log_{10} E = B_0 + B_1 \log_{10} S + B_2 \log_{10} N \qquad (3.3)$$

Assuming this equation describes the E,N,S relation, the unknown constants can be found by fitting a least squares surface to the points in the maximum error matrix. This was done using a Multiple Regression Analysis Routine on a

Control Data G-20 computer. For the particular problem being considered these coefficients were found to be.

$$B_0 = 1.936$$

$$B_1 = -0.131$$

$$B_2 = -0.889$$

The equation of the three-dimensional surface is therefore

$$\log_{10} E = 1.936 - 0.131 \log_{10} S - 0.889 \log_{10} N \qquad (3.1)$$

or

$$E = (86.3)\, S^{-0.131}\, N^{-0.889} \qquad (3.3)$$

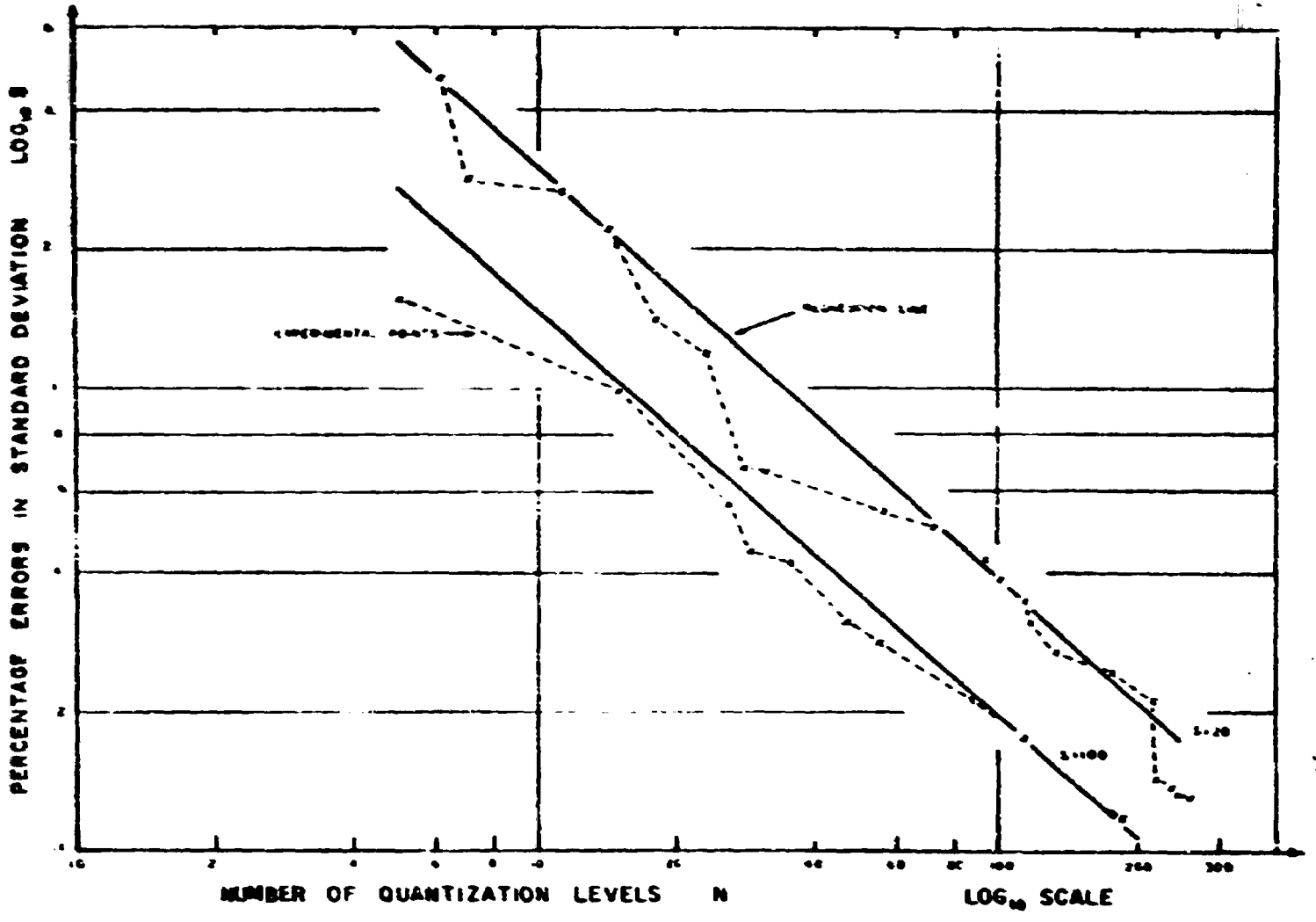
Using these equations it is possible to obtain the maximum errors in the standard deviation as a function of the sampling rate and the number of quantization levels. By allowing one of the terms in the above equation to be constant while the other two terms vary, it is easy to obtain the various cross sections of the surface. Typical cross sections have been plotted in Figures 4.2, 4.4 and 4.6. These curves will be examined and discussed in the next section.

4. A DISCUSSION OF THE OBSERVATIONS

It is important to first examine Figure 4.1 which gives a comparison of the smoothed curves obtained from the regression analysis routine, and the curves plotted using the raw data given by the matrix of the ordered maximum errors, (ordered with respect to increasing N). The results are encouraging because the deviation between the two curves is relatively small, and over a large portion of the range of N the regression lines lie above the actual maximum errors. As a result, the regression equation certainly yields worst-case results. The curves plotted in Figures 4.2, 4.4, and 4.6 are also worst-case since they were obtained from the same regression equation as that mentioned above. Therefore Figure 4.1 lends justification to the observations and conclusions made on the basis of the curves obtained using the regression coefficients.

The first of these curves is illustrated in Figure 4.2 which demonstrates the effect which an increased sampling rate has in reducing the percentage errors for a range of values of quantization levels. In all cases increasing the sampling rate significantly reduced the errors in the standard deviation. A qualitative explanation of this result will be given in the following paragraph.

As has been pointed out previously in connection with the Sampling Theorem, the statistics (at least those based on the first and second moments) of any waveform are defined by its power spectrum. Sampling the waveform in time can be looked on as using portions of the area under the power spectrum to compute the statistics. If the sampling frequency is too low, not all of the area under the power spectrum is utilized and the resulting values of the statistics are


PERCENTAGE ERRORS IN STANDARD DEVIATION
NUMBER OF QUANTIZATION LEVELS N
LOG SCALE

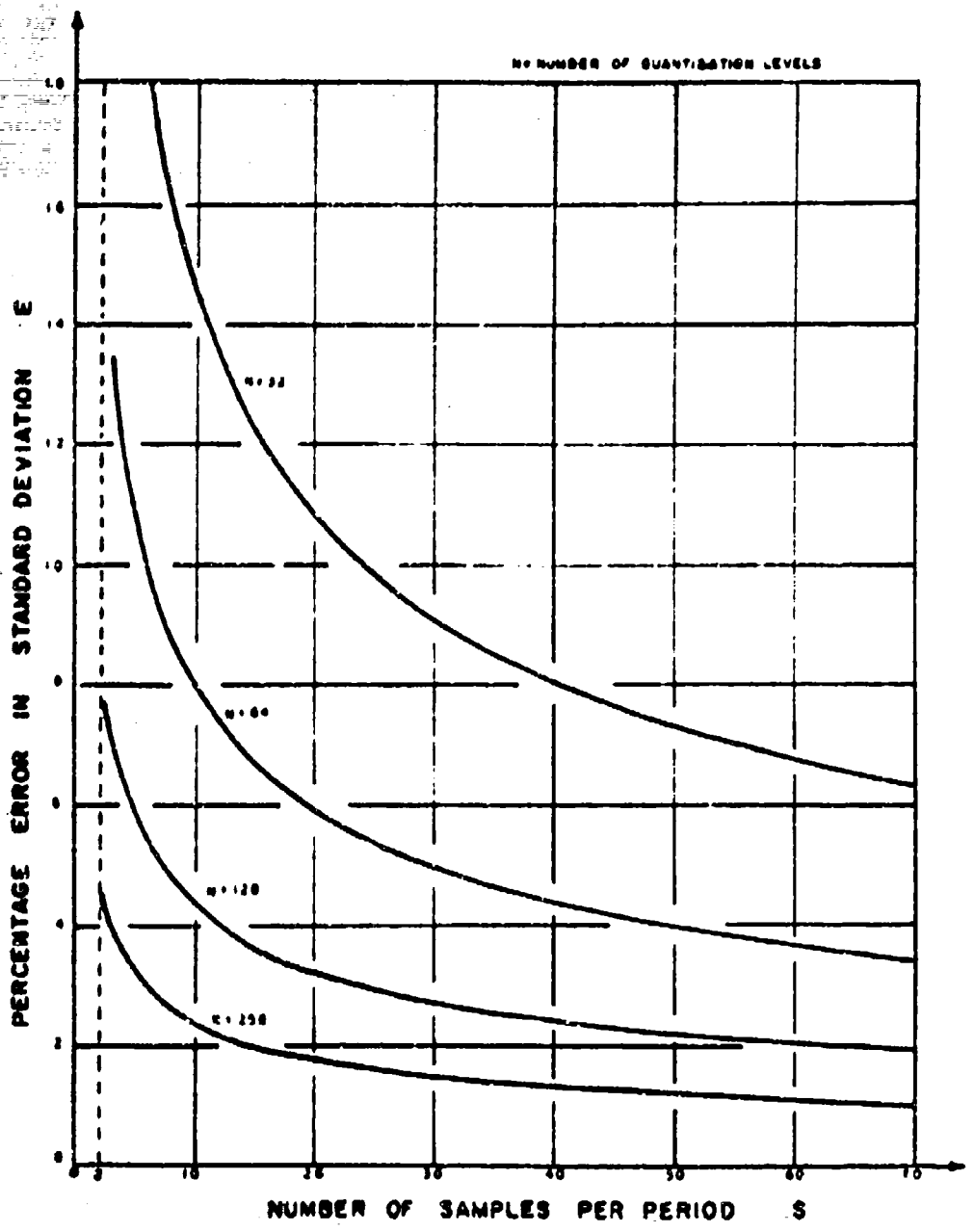

N= NUMBER OF QUANTISATION LEVELS
PERCENTAGE ERROR IN STANDARD DEVIATION E
NUMBER OF SAMPLES PER PERIOD S
10
20
30
40
50
60
70

Figure 4.2

in error. If the sampling frequency is high enough all of the area is accounted for and the statistics of the sampled and unsampled waveforms are the same. As shown in Figure 4.3a the power spectrum for the sine wave consists of two "spikes" at frequencies f_m and $-f_m$. By sampling at frequencies greater than or equal to $2f_m$ all of the power spectrum is accounted for and the statistics of the sampled and unsampled sine waves are the same.

It is well-known[1] that quantizing the sampled amplitudes adds noise to the waveform. If, for example, this noise is assumed to be white and independent of the sine waves, then the power spectrum for the digitized waveform looks like that shown in Figure 4.3b. If the statistics for this digitized waveform are to be the same as those for the original sine wave, then the sampling frequency would have to be infinitely high. It is sufficient to note that as the sampling frequency is increased to values greater than $2f_m$ more of the area under the power spectrum will be considered and the statistics will approach more closely their true values. This accounts for the shape of the curves plotted in Figure 4.2 which illustrate the tendency of the errors to be smaller as the sampling rate is increased. Further considerations in this regard can be found in the literature[3-6].

Another interesting result is brought out by Figure 4.1. These curves illustrate the fact that as the number of quantization levels increases the percentage error in the standard deviation decreases. This result can be explained using the concept of probability since it is possible to consider the perturbation of the amplitude from its sample value due to the quantization process as a random fluctuation. This deviation of the amplitude from its true value is due to what has been termed quantization noise[1].

The amplitude of this noise component ranges over the quantizing interval ΔV and can have any value inside this interval with equal probability. Figure

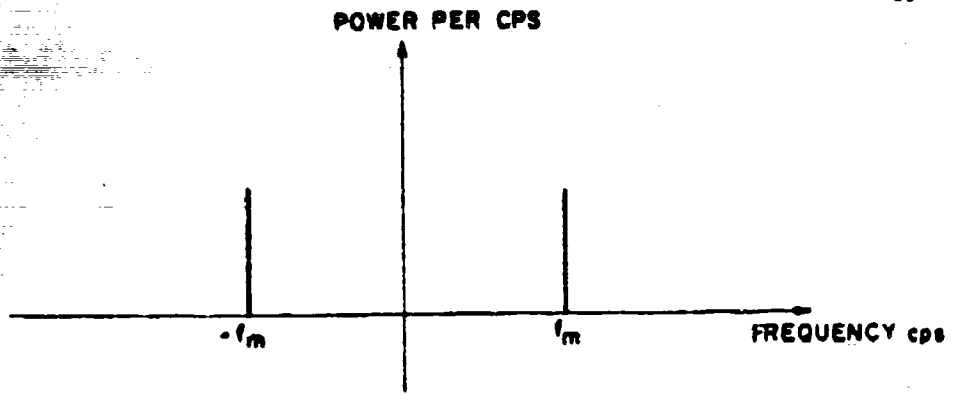


Figure 4.3a Power spectrum for a sine wave

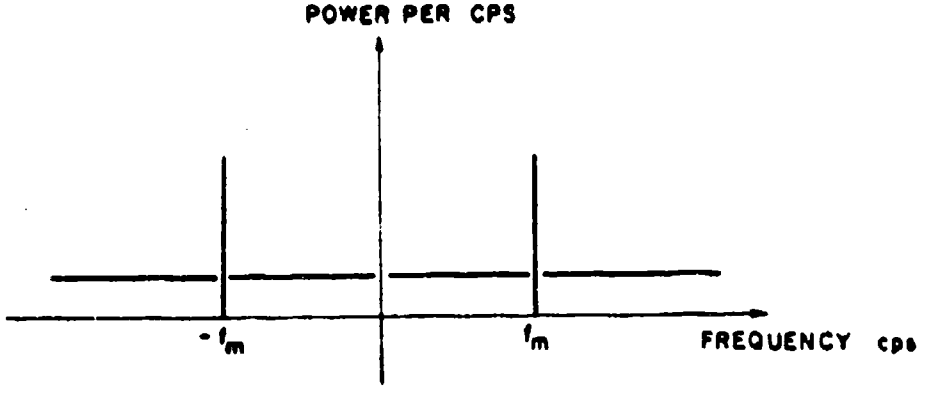


Figure 4.3b Power spectrum for the digitized sine wave (sine wave plus noise)

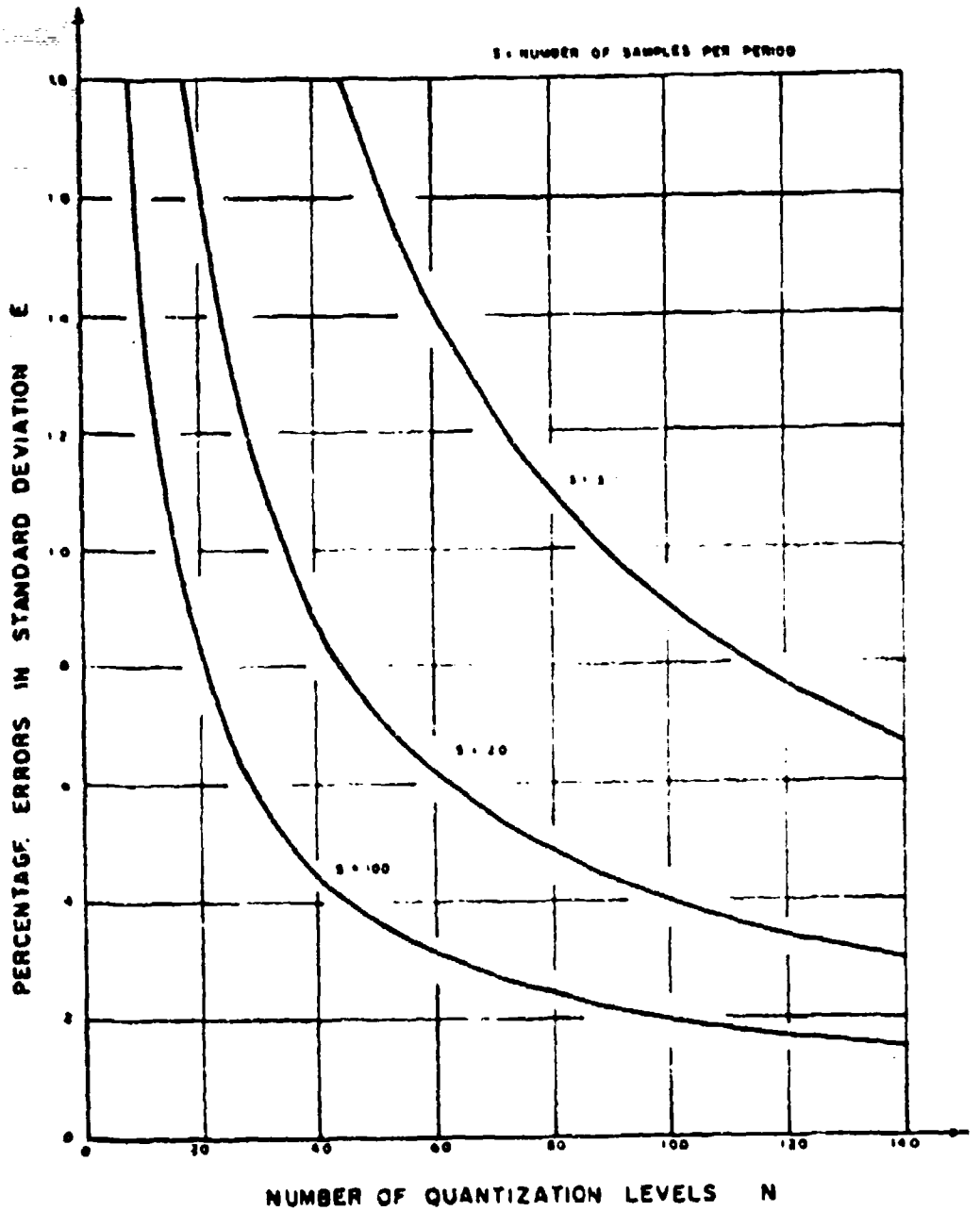

S : NUMBER OF SAMPLES PER PERIOD
PERCENTAGE ERRORS IN STANDARD DEVIATION E
S = 1
S = 40
S = 100
30
40
60
80
100
120
140
NUMBER OF QUANTIZATION LEVELS N

Figure 4.1

4.5 illustrates the behavior of a sample of quantization noise. The error is given by the difference between the amplitude of the actual waveform and the quantized result. Obviously this error ranges from $-\frac{\Delta V}{2}$ to $-\frac{\Delta V}{2}$. The average error will be zero. The second statistical moment is given by

$$E[\epsilon^2] = \int_{-\frac{\Delta V}{2}}^{\frac{\Delta V}{2}} \epsilon^2 \, p(\epsilon) \, d\epsilon = \frac{1}{\Delta V} \frac{\epsilon^3}{3} \Bigg]_{-\frac{\Delta V}{2}}^{\frac{\Delta V}{2}} = \frac{\Delta V^2}{12}$$

Since the mean error is zero, the RMS error is simply

$$\sqrt{\frac{\Delta V^2}{12}} = \frac{\Delta V}{\sqrt{12}}$$

Using the definition of the quantizing interval given in the preceding pages, $\Delta V = \frac{1}{N}$ where N is the number of quantization levels. Therefore increasing the number of quantization levels reduces the effects of the quantization noise since the error contribution is correspondingly smaller, as demonstrated by the result just derived.

The final observation which can be made using these results makes use of the curves plotted in Figure 4.6. Actually these curves contain the same information as the curves already discussed, but in a form which has more practical value. They illustrate the compromise or "trade off" which is possible between the number of quantization levels and the sampling rate to insure that the percentage error is no more than a certain specified value. For example, if it were desirable to set the quantizer at N = 128 and have errors less than 1.2%, the signal would have to be sampled at a rate S = 7 samples per period. For errors less than 1.1% the sampling rate would have to be increased to S = 31 samples per period. For a

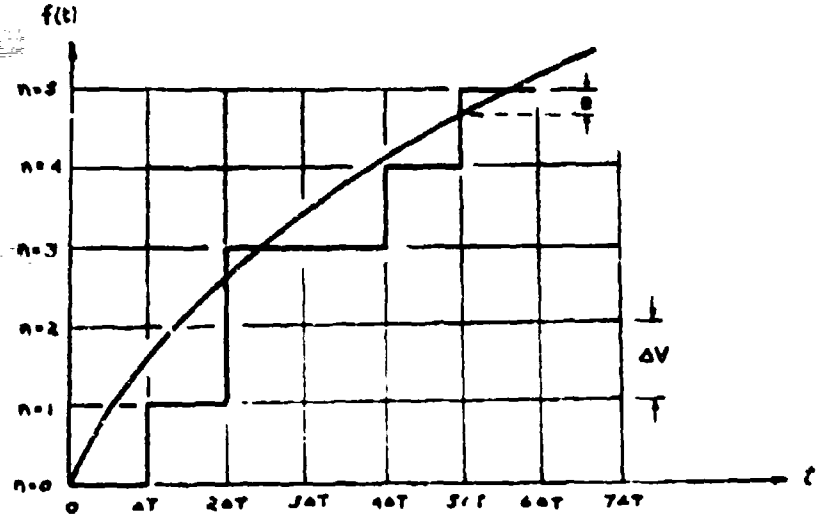


Figure 1.5a A typical error due to amplitude quantization

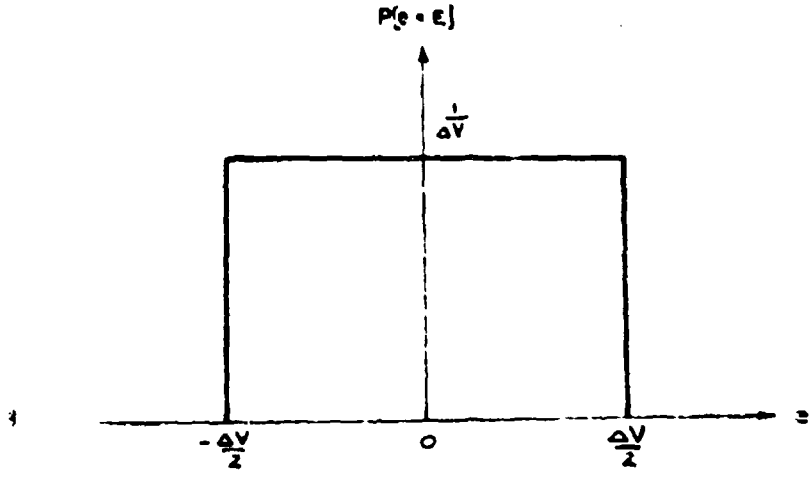


Figure 1.5b Probability density function of error due to amplitude quantization

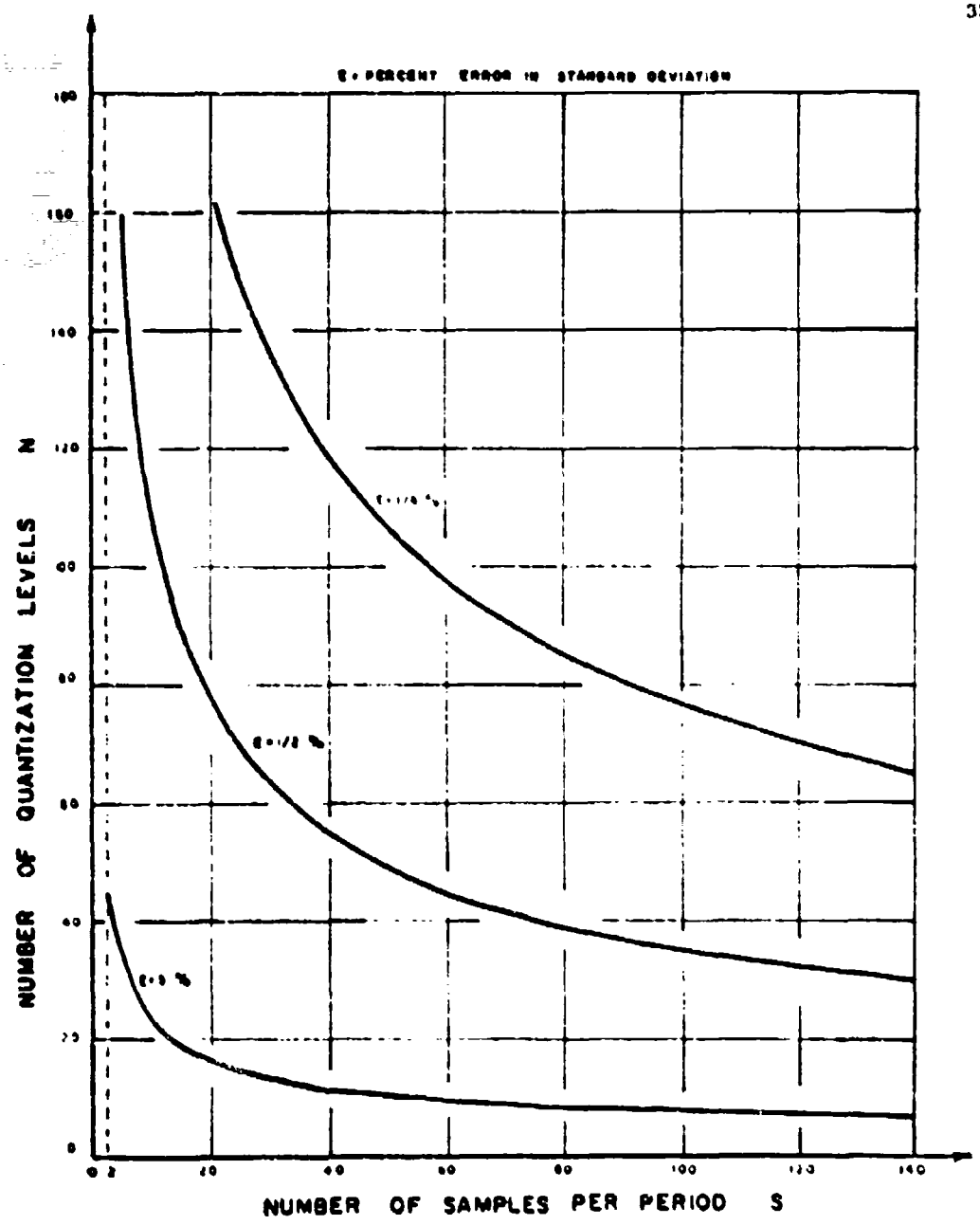

NUMBER OF QUANTIZATION LEVELS N
NUMBER OF SAMPLES PER PERIOD S

Figure 1.6

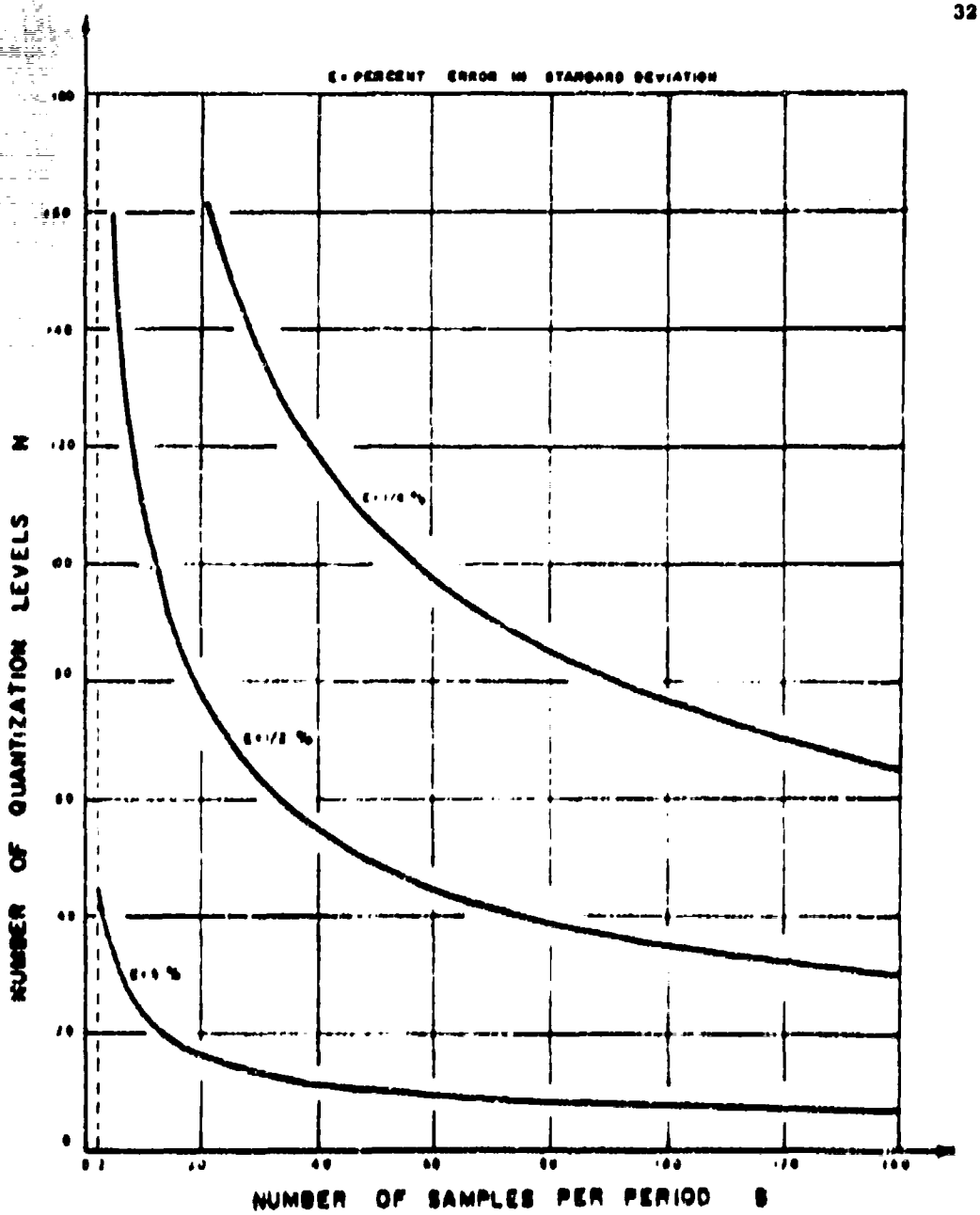

E = PERCENT ERROR IN STANDARD DEVIATION
NUMBER OF QUANTIZATION LEVELS N
NUMBER OF SAMPLES PER PERIOD S

Figure 1.4

smaller number of quantization levels the corresponding sampling rates have to be increased considerably. For N = 64 levels, then for an error less than 1.2%, S = 29 compared to S = 7 in the previous case. For an error less than 1.1% S = 111 must be used whereas S = 31 was satisfactory when N = 128.

5. CONCLUSIONS

The curves of Figure 4.4, which have demonstrated how the percentage errors in the standard deviation between the perturbed and unperturbed waveforms were reduced when more quantization levels were used, have been explained by treating the quantization process as one which adds a noise component to the sampled amplitudes. The advantage of this method of analysis is obvious when quantitative rather than qualitative explanations are desired. When noise is the subject of a statistical analysis, the results improve as the number of samples increases, (that is, as the number of quantization levels increases). Therefore, if one is interested in the formulation of a theoretical basis for the sampling and quantizing processes, it is the opinion of the author that this formulation should be based on the notions of probability and statistical theory. For further investigation using this approach interested readers are referred to the paper by Watts[2] which gives a fairly general treatment as to how one might handle statistical quantities when the processes on which they are defined are subjected to both sampling and quantization.

Throughout this paper an attempt was made, whenever possible, to point out useful statistical relations to suggest the possible application of statistical theory. In fact the curves drawn in Figures 4.2, 4.4, and 4.6 are expressed in terms of the sample standard deviation, itself a basic statistical quantity.

The conclusions which can be drawn on the basis of the preceding work are the following:

1. If a sine wave is to be sampled and quantized the relation between the number of quantization levels and the sampling rate which must be used to produce errors in the sample standard deviation less than a specified value is given by the curves of Figure 4.6, where each curve is plotted for a constant value of

this error.

2. The application of the quantization process to the sampled waveform has the effect of adding noise to that waveform. The effects of this noise are reduced by increasing the number of quantization levels or by increasing the sampling rate to a value in excess of that specified by the Sampling Theorem. As a result it is possible to reduce the errors in the values of the statistics introduced by quantization by increasing the sampling rate beyond its lower bound.

3. It seems that the most fruitful approach to the theoretical study of this problem would be to extend the work done by Watts[2] to the case where the bounds on the sampling rate and number of quantization levels become parameters in the expressions for the probability functions. The statistics could then be evaluated as functions of these parameters and compared directly with the true values. This would be an extremely useful result since the expressions for the digitized waveform depend primarily on the probability functions of the unperturbed waveform. The results would then be applicable to all processes for which a probability distribution could be found.

BIBLIOGRAPHY

1. C. E. Shannon, "Communication in the Presence of Noise," I.R.E. Proc., Vol. 37, January 1949, pp 10-21.

2. D. G. Watts, "A General Theory of Amplitude Quantization with Applications to Correlation Determination," Proc. I.E.E., Vol. 100, Part C, March 1962, pp 209-218.

3. H. D. Brunck, An Introduction to Probability and Statistics, Ginn and Company, New York, 1962, pp 48-49.

4. B. Widrow, "A Study of Rough Amplitude Quantization by Means of Nyquist Sampling Theory," I.R.E. Trans., 1956, CT-3, pp 266-276.

5. H. D. Holms and F. B. Thomas, "Truncation Error of Sampling Theorem Expansions," I.R.E. Proc., Vol. 50, February 1962, pp 179-184.

6. Stewart, "Statistical Design and Evaluation of Filters for the Restoration of Sampled Data," I.R.E. Proc., February 1956, pp 253-257.

7. M. Schwartz, Information Transmission Modulation and Noise, McGraw-Hill, New York, 1959.